



Publishing Species Checklists Complete Guide

Version 2.0



April 2011



Suggested citation:

GBIF (2011). Publishing Species Checklists, (contributed by Remsen D., Döring M., Robertson, T., Ko, B.), Copenhagen: Global Biodiversity Information Facility, 53 pp, accessible online at http://links.gbif.org/checklist_how_to

This document provides instructions on sharing species checklists using an international data exchange format known as Darwin Core Archives.

ISBN: 87-92020-32-1

Persistent URI: http://links.gbif.org/checklist_how_to

Language: English

Copyright © Global Biodiversity Information Facility, 2010

License:

This document is licensed under a [Creative Commons Attribution 3.0 Unported License](#)

Document Control:

Version	Description	Date of release	Author(s)
1.0	Release Version	1 April 2011	David Remsen
1.01	Minor revisions	15 April 2011	David Remsen
2.0	Major update and consolidation	22 August 2011	David Remsen

Cover Art Credit: Gregory Basco
Yellow-throated Euphonia, *Euphonia hirundinacea*

GBIF Contact

David Remsen (dremsen@gbif.org) regarding the composition and format of this document and its source files.

This document is also part of the 'GBIF Data Publishing Manual version 1.0, ISBN 87-92020-31-3, available at http://links.gbif.org/data_publishing_manual

About GBIF

The Global Biodiversity Information Facility (GBIF) was established as a global mega-science initiative to address one of the great challenges of the 21st century - harnessing knowledge of the Earth's biological diversity. GBIF envisions 'a world in which biodiversity information is freely and universally available for science, society, and a sustainable future'. GBIF's mission is to be the foremost global resource for biodiversity information, and engender smart solutions for environmental and human well-being¹. To achieve this mission, GBIF encourages a wide variety of data publishers across the globe to discover and publish data through its network.

¹ GBIF (2011). GBIF Strategic Plan 2012-16: Seizing the future. Copenhagen: Global Biodiversity Information Facility. 7pp. ISBN: 87-92020-18-6. Accessible at http://links.gbif.org/sp2012_2016.pdf

Table of Contents:

About GBIF	3
Introduction	5
Scope: What is meant by “Checklist”	5
The Darwin Core Archive Format	7
Checklist Data: Core and Extension files	7
Checklist Metadata documents	8
Composing a Checklist Metadata document	9
The Data Descriptor File.....	9
Validating and Publishing a Darwin Core Archive species checklist.	11
Best Practices in Publishing Species checklists using the GNA Profile.....	12
Sharing Scientific Names	12
Publishing Classifications.....	14
Classification Formats <u>not</u> recommended for publishing.....	17
Publishing Synonymy	18
Citation and Attribution	21
Authenticated File Access via httpd.....	24
The Core Taxon Class	26
Sharing Vernacular Names	32
Vernacular Names Extension.....	32
Sharing Species Descriptions	35
Species Description Extension	35
Sharing Species Distributions	37
Species Distribution Extension.....	37
Sharing References.....	40
References Extension	40
Sharing Type information	43
Types and Specimens Extension.....	43
Sharing Links and Identifiers	46
Creating a dynamic link to a species page.....	46
Alternative Identifiers extension.....	47
Resource Relationship Extension	49
Summary: Abbreviated listing of the GNA Profile.....	51

Introduction

This guide provides details on how to utilise the Darwin Core Archive (DwC-A) format as a means to share taxonomic checklist information in a standard way using a specific set of standardized terms called the Global Names Architecture (GNA) Profile. This guide provides an overview of these terms, the data format, and recommended practices to be used when sharing species checklist information.

The DwC-A format and the GNA profile described here represent an internationally recognised and ratified data exchange format for sharing taxonomic data. The Darwin Core Archive format represents an intermediate position between the two ends of a continuum of data complexity. All data exchange standards strike a balance between complexity and coverage with simple solutions sacrificing coverage and complexity in favour of ease-of-use. Highly complex formats provide more complete solutions for representing any type of data but at the expense of simplicity and require supporting software and expertise. The GNA profile focuses on the key elements of taxonomic checklists and enables an enriched set of data types to be linked to this core structure. The data contained in an archive can be readily understood and used by many biologists and data managers familiar with basic structured text files. By providing an international standard that is relatively easy to produce and consume, and that supports many of the key elements that compose a taxonomic data resource, GBIF hopes to provide the creators and managers of checklists with a standardised approach to sharing their data and promote common approaches to the subsequent citation and recognition of their work. A standard format also increases relevance and utility.

Scope: What is meant by “Checklist”

The terms “species checklist” and taxonomic “catalogue” may refer to an overlapping range of taxonomic resources. All of these products contain sets of scientific names that implicitly or explicitly refer to taxa. The set of names included in such a list may be constrained by taxonomic group, geographic region, or by a theme, such as invasive species, or some combination of all three. In order of increasing detail these include the following resource types²

² These categories and descriptions are derived directly from “Hyam . R., Standardisation of Data Exchange in the Pan-European Species-directories Infrastructure (PESI) Deliverable D 4.1”

1. **Name lists** - Simple lists of species names with no explicit indication of taxonomic status, but generally implied to consist of accepted names of taxa. Such lists are generally intended to identify a set of taxa included within some regional or thematic context.
2. **Nomenclatural lists (Nomenclators)** - Lists of names including the nominal taxa, meaning the registry of published usages of scientific names representing nomenclatural acts as governed by the respective Codes of Nomenclature. Most of these acts are ‘original descriptions’ of new scientific names, but other acts may include emendations, lectotypifications, and other acts as governed by the Codes. Synonymy is not included in these lists as taxonomic concept, but only as newly established combination (for botanists) linked to a basionym, thus providing a nomenclatural synonym.
3. **Taxonomic checklist** - These lists extend nomenclatural lists by adding taxonomic opinion in the form of explicit taxonomic status information and the inclusion of names placed in synonymy. Simple taxonomic lists in this category provide no specific circumscription details regarding the basis for the synonymy. Taxa are often organised into classifications. The term “*taxonomic catalogue*” may also be used to refer to instances of this, and the remaining categories, particularly if the resource includes verified publication and taxonomic status details.
4. **Annotated Checklists** - This category extends taxonomic checklists by adding other data types (annotations) to the core, synonymised checklist, such as common names, threat status, distribution and basic descriptive information. When the annotation types provide sufficient detail to effectively define, or circumscribe, a taxon, such as detailed diagnostic descriptions and illustrations, molecular data, specimens, etc., then the annotated list may fall into one of the two categories defined below.
5. **Flora or Faunal lists** - These are typically books that provide detailed species accounts for a particular region. Details may include many of the data types included in annotated checklists but include specific data types, such as detailed descriptions and illustrations, specimen references and other details that explicitly define (circumscribe) the taxon within the scope of the region which is not necessarily global.
6. **Monographs** - Monographs are detailed species accounts often published as books for a particular taxon group at the global scale. It will contain detailed nomenclature and synonymy and taxon circumscription details, that include textual descriptions and illustrations, details of specimens examined and included in the definition, and a bibliography of examined literature.

The Darwin Core Archive format with the GNA profile supports the exchange of key data elements within all of these checklist data types. The specific degree of coverage depends very much on the individual resource. In this document we will use the term ‘checklist’ in the broad sense as a general term for referring to any or all of the resource categories described above. The specific category types will be used when a specific resource is to be referenced.

The Darwin Core Archive Format

Darwin Core Archive (DwC-A) is an informatics data standard that makes use of the Darwin Core terms to produce a single, self-contained dataset for checklist data. The collection of files in an archive form a self-contained dataset, which can be provided as a single compressed (Zip or GZIP) file. An archive consists of the following files.



A required *core data file* containing the species checklist itself, classification and nomenclatural details. Data elements and sharing best practices are included in this manual.



A required *metadata file* that describes the checklist dataset and provides provenance information. Options for creating this file are included in this manual.



Optional *extension files* that describe additional data types, such as common names or species descriptions, that annotate the core species list. Data elements and sharing best practices are included in this manual.



A (nearly always) required *data descriptor file* that describes the data files in the archive. Options for generating this file are provided in this manual.



The collection of files is compressed into a single archive file that can be easily hosted, transferred and opened for use.

Checklist Data: Core and Extension files

The Darwin Core Archive format provides the structural framework for publishing species checklists. The data files, corresponding to the core and extension files, in a Darwin Core

Archives consist of one or more text files, in standard comma- or tab-delimited format. The files are logically arranged in a star-like manner with the *core file*, containing the basic checklist elements (species list, classification, synonymy) surrounded by any number of *extensions* files. Links between core and extension records are made using a taxon identifier (*taxonID*) data element. In this way, many extension records can exist for each single core taxon record. This “star-schema” provides a simple relational data model that supports many types of annotations that are common to species checklists.

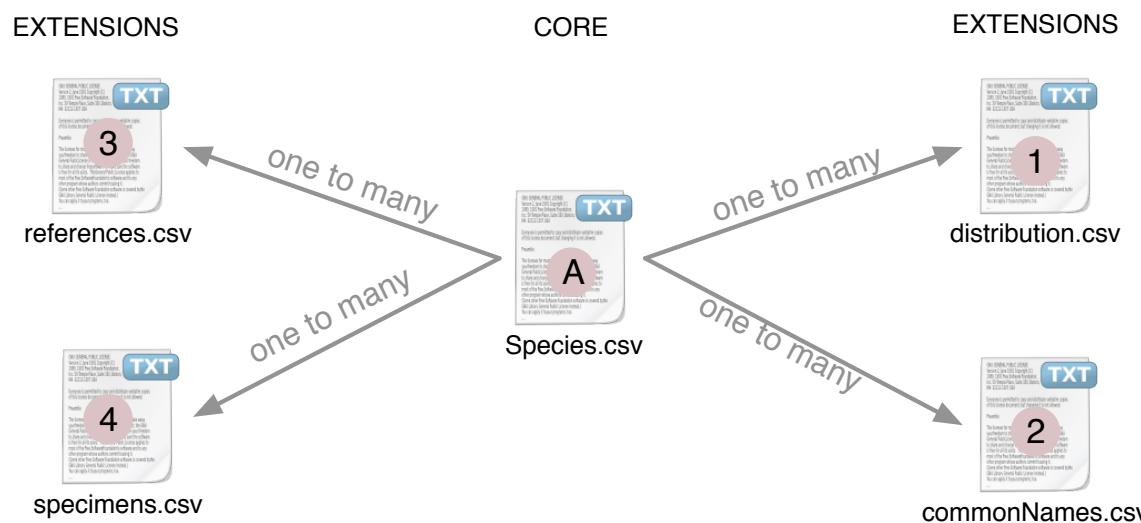


Figure 1 - Darwin Core Archive data files in 'star schema' linked by a common identifier, *taxonID*

Checklist Metadata documents

Documenting the provenance and scope of datasets is required in order to publish checklist data through the GBIF network. Dataset documentation is referred to as ‘resource metadata’ and enables users to evaluate the fitness-for-use of a dataset. It may describe the scope and intended function of the list, methodologies and resources used for its compilation, and the individuals and organisations involved in its creation and management. Given the broad scope of species checklists, a metadata document is considered to be mandatory for sharing checklists through the GBIF network. Metadata is shared in a Darwin Core Archive as an XML document. GBIF provides a metadata profile for species checklists based on the Ecological Metadata Language³.

³ Ecological Metadata Language (EML) <http://knb.ecoinformatics.org/software/eml/>

For more information, the GBIF Metadata Profile Reference Guide provides an overview of the data elements in the GBIF profile of the Ecological Metadata Language used to document a species checklist resource (abstract, contributors, attribution, etc.).

http://links.gbif.org/gbif_metadata_profile_en_v1

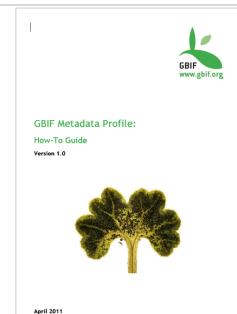


Composing a Checklist Metadata document

GBIF provides two primary mechanisms to create a compliant metadata document for checklists: through the Integrated Publishing Toolkit (IPT)⁴ and using a pre-configured Excel spreadsheet template in combination with the GBIF Spreadsheet web processor⁵ which converts a completed spreadsheet file to an EML document. The IPT provides a form-based interface for creating a metadata document and is the preferred method for composing and managing metadata when the IPT is intended to be installed and used for publishing species checklists. The spreadsheet-based solution is intended for use when the data publisher is planning on generating an archive directly (without the IPT).

The GBIF Metadata Profile, How-to Guide provide a more complete review of the two primary methods for composing a metadata document that conforms to the GBIF metadata profile: using the Integrated Publishing Toolkit or using a MS Excel Metadata template.

http://links.gbif.org/gbif_metadata_profile_how-to_en_v1



The Data Descriptor File

Darwin Core Archives require the inclusion of a special XML descriptor file that describes the fields contained in the various data files and their mappings to the GNA profile. It also contains physical settings for these files and other technical details.

There are two ways to generate this descriptor file for inclusion with species checklist archives.

⁴ Integrated Publishing Toolkit - <http://code.google.com/p/gbif-providertoolkit/>

⁵ GBIF Spreadsheet templates and processor - <http://tools.gbif.org/spreadsheet-processor/>

1. Use the Integrated Publishing Toolkit (IPT) to publish a species checklist. It will generate the descriptor file automatically and include it in the zipped archive.
2. Use the Darwin Core Archive Assistant⁶. This is a tool that provides a simple point-and-click interface to describe the data files in an archive. It generates the required XML that can be downloaded and saved.

Darwin Core Archives may be composed and published with no local tools installed beyond a basic web server. A guide, Create your own Darwin Core Archive, Step-by-Step Guide provides more complete details.

http://links.gbif.org/dwca_manual_create_en_v1

This guide provides step-by-step instructions on composing and hosting your own Darwin Core Archive. This method is a good solution for technically capable database administrators who may wish to develop an automated scripted solution for maintaining a regularly updated data archive. This solution also describes the Darwin Core Archive Assistant, a service for composing a valid XML descriptor file without having to manually edit XML.

Darwin Core Archive Assistant, User Guide

http://links.gbif.org/gbif_dwc-a_asst_en_v1.1



For manual creation of an archive, please keep the following in mind:

- It is recommended to use TAB or Comma-Separated-Values instead of custom field delimiters and quotes.
- Be careful and consistent with quotation.
- Encode text files as UTF-8
- Make sure you replace all line breaks in a data field, i.e. \r \n or \r\n with either simple spaces or use 2 characters like “\$\$” to replace "\r" to escape the line break if the intention is to preserve them. Another option is to replace line breaks with the html
 tag.
- Encode NULLs as empty strings, i.e. no characters between 2 delimiters, or \N or \NULL, but no other text sequence!

⁶ Darwin Core Archive Assistant - <http://tools.gbif.org/dwca-assistant/>

Darwin Core Archive, How-to Guide

http://links.gbif.org/gbif_dwc-a.How_to_guide_en_v1

A more general introduction to the Darwin Core Archive format is available and may be useful for those data administrators interested in automating the Darwin Core Archive creation process using database scripts.



Validating and Publishing a Darwin Core Archive species checklist.

Users publishing species checklists using the Integrated Publishing Toolkit (IPT) can skip this section. Users creating an archive manually should first validate their archive prior to publishing it through the GBIF network. This can be done using the Darwin Core Archive Validation Service.

<http://tools.gbif.org/dwca-validator/>

which checks the contents of the archive and the syntax of the descriptor file to ensure it conforms to the standard. The validator provides detailed feedback.

Once an archive conforms to the Darwin Core Archive standard, it can be published to the GBIF network. Users of the IPT have a publication system build into the tool which enabled publication at the push of a button. Manually generated archives can be published to the GBIF network in two simple steps.

1. The archive must be hosted on a publically accessible web address (URL).
2. The address and other basic information must be provided to GBIF using a simple registration form available at <http://tools.gbif.org/dwca-register/>

Best Practices in Publishing Species checklists using the GNA Profile

For ease in understanding, we may use the terms *field* in this guide to refer to the Darwin Core set of terms in the taxonomic publishing profile to which a users data will be mapped. For example, we will refer to the use of the *dwc:scientificName field* when referring to the Darwin Core term, *scientificName*.

Sharing Scientific Names

The Darwin Core supports more than one way to share a scientific name. This includes the following options:

A. Concatenated in the *scientificName* field

scientificName
Gerardia paupercula var. borealis (Pennell) Deam

The *dwc:scientificName* field stores the full scientific name of a taxon including authorship. This field should always be populated with data even if the names are split into component parts (as in C. below). Databases that do not provide a clean separation between the name part and the authorship part of the name should use this field for the entire concatenated name string. This may be needed for hybrid formulas, *sensu strictu* names, autonyms and other non-trivial binomials. This field is generally used in combination with the *dwc:taxonRank* field to store the scientific name parts of a full taxonomic list including the higher taxa.

B. Separate Name and Authorship parts

scientificName	scientificNameAuthorship
Gerardia paupercula var. borealis	(Pennell) Deam

Some databases separate a scientific name into a name part and an authorship part. In this case the *dwc:scientificName* and *dwc:scientificNameAuthorship* fields should be used.

C. Separated into name parts.

Genus	specificEpithet	taxonRank	infraspecificEpithet	scientificNameAuthorship
Gerardia	paupercula	var.	borealis	(Pennell) Deam

The Darwin Core provides a series of terms that enable scientific names to be separated into component parts. Some databases store species lists in such parsed components. In

this case, sharing data in this form may be an option. If so, however, it is strongly recommended that an additional and complete name be composed from the parts and shared in the *dwc:scientificName* field (as in section A above). Note that in the table above, the Darwin Core term, *dwc:subgenus*, is not displayed but represents an additional name component.

Infrageneric Markers

If possible, please provide an infrageneric rank marker as part of the scientific name to avoid confusion with the original / basionym author. For example “*Ageratina subgen. Apoda* R.M.King & H.Rob” is preferred over “*Ageratina (Apoda)* R.M.King & H.Rob.” as the later *Apoda* could interpreted as a subgenus or as the basionym author.

Publishing Classifications

The Darwin Core provides two basic options for publishing classifications or taxonomic hierarchies; normalized and denormalised. These two options account for the primary means by which most classifications are managed in databases.

Normalised Classifications (Parent/Child)

The recommended way to share a classification is in a normalized format. This may also be referred to in a database as a "parent-child relationship" or an "adjacency list". In a normalized taxonomic hierarchy, each taxon is represented by a single row. This includes both species and all higher taxa in the classification. Each row has at least the following component data elements.

- A *dwc:taxonID* referring to the current taxon.
- The *dwc:scientificName* of the current taxon. Example: "*Panthera tigris*"
- The *dwc:taxonRank* of the referent taxon. Example: "*species*"
- A reference to the taxon identifier of the immediate parent taxon stored in the *dwc:parentNameUsageID*. In the example below, the parent of record 7, for "*Panthera tigris* (Linnaeus)" is record 6, the genus "*Panthera*."

A sample classification for a single species, the tiger, "*Panthera tigris*", is illustrated below. Note that the top-most member of a hierarchy has no parent so that the parent identifier should be null or a "0". Note that *dwc:scientificName* provides a common field for storing the name in this case but that the full set of options for names is described above in Sharing Scientific Names.

taxonID	taxonRank	scientificName	parentNameUsageID
1	Kingdom	Animalia	0
2	Phylum	Chordata	1
3	Class	Mammalia	2
4	Order	Carnivora	3
5	Family	Felidae	4
6	Genus	Panthera	5
7	Species	Panthera tigris (Linnaeus)	6

Advantages

- *Efficiency* - A normalised classification stores a single reference for each taxon in the hierarchy.
- *Referential integrity* - Each taxon component has a distinct identifier that explicitly references its immediate parent. It is easy to verify that the taxonomic hierarchy is complete and properly formed.
- *Extensibility* - All taxa are identified with distinct taxon identifiers. This enables higher taxa to be more richly documented through the use of extensions in the same manner as species records.

Disadvantages

- *Convenience* - A normalised classification does not provide an intuitive view of the classification hierarchy when viewed in raw tabular form. Many biologists manage classifications in a less efficient, but more visually intuitive, *de-normalised format*, described below. Transforming a de-normalised classification to the normalized form is difficult to manually perform.

Note: A *dwc:parentNameUsageID* must point to an existing record in the dataset. It is invalid to point to higher taxon identifiers that do not exist as records.

Denormalised Classifications

This format is familiar to anyone who manages species information in spreadsheets. In a de-normalised classification, each row of the data table refers to one of the terminal taxa, such as a species, and a complete set of parent taxa as a set of columns, one for each parent taxon.

This format is not the recommended method for sharing taxonomic data using Darwin Core Archives but is supported by GBIF as it is in common use in many species lists. If this is the method by which data will be shared, it is highly recommended that

1. Each higher taxon column is completely populated. Avoid blanks as in the Plantae example below.
2. Ensure taxonomic integrity of the list. For example ensure that two species in a common genus share the same family. Ensure that if synonyms are included in separate rows, that their classification matches that of the accepted taxon.

taxonID	kingdom	phylum	class	order	family	scientificName*
1001	Animalia	Chordata	Mammalia	Carnivora	Felidae	Panthera tigris
1002	Animalia	Chordata	Mammalia	Carnivora	Felidae	Panthera leo

1003	Animalia	Arthropoda	Insecta	Hymenoptera	Apidae	Apis mellifera
1004	Plantae	--	--	--	Poales	Poa annularis

Advantages

- *Legibility* - The primary advantage of this format is that it is easy to read and the taxonomic hierarchy can be evaluated by simply reading columns.
- *Convenient* - Spreadsheet applications and many relational databases make it easy to implement this structure for storing hierarchical data.

Disadvantages

- *Higher likelihood of referential integrity loss* - Higher taxa are repeated in this format which can increase the chance that two identical taxa may be spelled differently. Other similar risks are possible with this format. For example it is possible for two instances of the same taxon (example “Felidae”) to be assigned to two different parents, resulting in a conflict of hierarchical integrity.
- *Lack of details for higher taxa* - This format treats higher taxa as properties of a species, not as separate taxon records themselves. Therefore, this format does not allow properties of higher taxa to be shared either in the core file or in any extensions.

Other classification-related recommendations

- Try to include a Kingdom and a nomenclatural code reference for all records even for basic species lists.
- Try to include Kingdom, Phylum and Family as a minimal classification for de-normalised classifications.
- If it is the same throughout the dataset, consider using a static mapping of the term and value. See the Darwin Core Archive How-to Guide at http://links.gbif.org/gbif_dwca_how_to_guide_en_v1 for details on mapping global values.

Classification Formats not recommended for publishing

The following examples illustrate data configurations that can fit the profile but are not recommended or supported by GBIF (i.e., GBIF parsers would not handle these cases properly)

A. This example identifies the referent taxon as the last column containing taxon values.

taxonID	kingdom	phylum	class	order	family	scientificName*
997	Animalia					
998	Animalia	Chordata				
999	Animalia	Chordata	Mammalia			
1000	Animalia	Chordata	Mammalia	Carnivora		
1001	Animalia	Chordata	Mammalia	Carnivora	Felidae	
1002	Animalia	Chordata	Mammalia	Carnivora	Felidae	Panthera tigris
1003	Animalia	Chordata	Mammalia	Carnivora	Felidae	Panthera tigris

B. This example attempts is similar to A above but attempts to reduce integrity errors by only recording higher taxon names once.

taxonID	kingdom	phylum	class	order	family	scientificName*
997	Animalia					
998		Chordata				
999			Mammalia			
1000				Carnivora		
1001					Felidae	
1002						Panthera tigris
1003						Panthera leo

Please avoid publishing data in these configurations.

Publishing Synonymy

Darwin Core Archive supports the publication of synonyms in species checklists. A synonym is published as a separate record in the core data file. A synonym references the accepted taxon record through the use of the *dwc:acceptedNameUsageID* field. This field contains the *dwc:taxonID* representing the accepted taxon record. In the simplified example below, the first record represents the accepted name for a taxon and records 2 and 3 are synonyms.

taxonID	scientificName	acceptedNameID	taxonomicStatus	nomenclaturalStatus
1	Coeligena helianthea (Lesson 1838)	1	accepted	
2	Ornismya helianthea Lesson 1838	1	Homotypic synonym	
3	Helianthea helianthea (Lesson 1838) J. Gould 1848	1	Homotypic synonym	
4	Helianthea typica Bonaparte 1850	1	Heterotypic synonym	nomen dubium
5	Helianthea porphyrogaster Mulsant 1876	1	Heterotypic synonym	nomen dubium
6	Coeligena helianthea tamai Berlioz & Phelps 1953	1	Heterotypic synonym	nomen dubium

A synonym record is recommended to contain a distinct *dwc:taxonID* or it may have no *dwc:taxonID* at all. It *must not* use the same *dwc:taxonID* as the accepted taxon record. The simplest representation of synonymy is as provided in the example above where synonyms are listed as distinct records and ‘point’ to the accepted taxon record using the *dwc:acceptedNameUsageID*. This simple synonymy supports the publication of basic taxonomic checklists with synonym details limited to the core taxon class elements. The *dwc:taxonomicStatus* field affirms the status of the record. A recommended vocabulary for this field is [available](#)⁷. Additional nomenclatural details that may also support the

⁷ http://rs.gbif.org/vocabulary/gbif/taxonomic_status.xml

rationale behind the synonymy may be included using the *dwc:nomenclaturalStatus* field and [supporting vocabulary](#)⁸.

Detailed synonymy can be supported by ensuring a unique *dwc:taxonID* is included in each synonym record and by utilising the available extensions to support the sharing of checklist annotations. This supports the linking of one or more bibliographic records, specimen records and other data types supported by the GNA Profile to a single synonym record in the core data file. If a *dwc:taxonID* is not provided for a synonym record, extensions cannot be used as they rely on the *dwc:taxonID* to provide the link to the taxon record in the core file. A simplified example below illustrates the use of two files (expressed as tables) to provide a bibliography for a synonym using the References extension. The shared *dwc:taxonID* is highlighted in the example.

Taxon.txt data file

taxonID	scientificName	acceptedNameUsageId	taxonomicStatus
1	Coeligena helianthea	1	accepted
2	Ornismya helianthea	1	synonym
3	Helianthea helianthea	1	synonym

References.txt data file

taxonID	Bibliographic citation
2	Schmidt, O. 1870. Grundzüge einer Spongiens-Fauna des atlantischen Gebietes. (Wilhelm Engelmann: Leipzig): iii-iv, 1-88, pls I-VI.
2	Laubenfels, M.W. De 1942. Porifera from Greenland and Baffinland collected by Capt. Robert A. Bartlett. Journal of the Washington Academy of Sciences 32(9): 263-269.

Other Synonymy Do's and Don'ts

- An *dwc:acceptedNameUsageId* must point to an existing record in the dataset. It is invalid to point to accepted taxa that do not exist as records.
- Do not confuse the *dwc:higherTaxonID* used to describe a classification with the *dwc:acceptedNameUsageId* used to describe the taxonomic status of a record.
- Do not “chain” synonyms. A synonym should only point to accepted taxon records via *dwc:acceptedNameUsageId*. They should never point to another synonym.

⁸ http://rs.gbif.org/vocabulary/gbif/nomenclatural_status.xml

Nomenclatural Synonymy

Nomenclatural synonymy is supported in the core data file through the use of the *dwc:originalNameUsageID* field. This field refers to the row representing the original taxon reference for the name. This record is recommended to provide a bibliographic citation in the *dwc:namePublishedIn* field, which refers to the publication in which the name was originally established.

taxonID	scientificName	originalNameID	namePublishedIn
1	Tetrao afer Müller 1778	1	J. Syst. Nat 7:31
2	Pternistes afer (Müller 1778)	1	
3	Francolinus afer afer (Müller 1778)	1	

Nomenclatural and taxonomic synonyms may be designated in the same taxon record.

Note: An *dwc:originalNameUsageID* must point to an existing record in the dataset. It is invalid to point to accepted taxa that do not exist as records.

Pro-partة Synonymy

Sometimes the same name may be a synonym for more than one accepted taxon or may be both an accepted taxon name and a synonym. These are caused by splits and circumscription changes where, for example, a series of types may be divided among multiple taxa. The recommended practice for sharing pro-partة synonyms is represented in the example. In this example, *Vireo solitarius* is an accepted taxon name and it is also included in the synonymy for both *Vireo cassinii* and *Vireo plumbeus*. In the case of the synonyms, they are represented as a single record with accepted taxon reference concatenated in the *dwc:acceptedNameUsageID* field and separated by a pipe (“|”) character.

taxonID	scientificName	acceptedNameUsageID	taxonomicStatus
1	<i>Vireo solitarius</i>	1	accepted
2	<i>Vireo cassinii</i>	2	accepted
3	<i>Vireo plumbeus</i>	3	accepted
4	<i>Vireo solitarius</i>	2 3	pro-partة

Citation and Attribution

Taxonomic checklists often represent significant intellectual and financial efforts on the part of the individuals and organisations who compile them. Some checklists may be derived from, or may reference, other source checklists to create new distinct thematic, regional or taxonomic views of the same source authority. Proper attribution and visibility of these sources is therefore a high priority.

The DwC-A format provides a range of options and recommendations for providing proper citation and attribution. This range extends from global citation and attribution information that form part of the resource metadata down to record-level data elements. These options support the provision of multiple levels of attribution.

Metadata Citation and Attribution

The GBIF Metadata profile supports resource-level data elements that contribute to citation and attribution and enable detailed description of the scope and provenance of a checklist. A complete reference list to all the metadata elements is beyond the scope of this document and [available](#)⁹ but specific citation and attribution-related elements include:

- **Intellectual Property Rights** - The metadata profile contains a rights management statement for the resource, or a reference to a service providing such information, such as a Creative Commons license. It also includes an element describing the intended use and purpose of the dataset.
- **Individuals and Organisations** - The metadata profile enables the description of any and all individuals, institutions or organisations that may be associated with a dataset. These agents may be ascribed different roles relative to the dataset and may include URLs to each resource. This section provides one method for describing and linking to individuals and organisations that have contributed to a checklist.
- **Source URL** - Links to the homepage of the source
- **Project Information** - If the checklist is linked to a particular project (e.g., “The Catalogue of Life”) there are a set of fields for describing the project in detail.
- **Citation** - This element allows the checklist publisher to specify exactly how the checklist data should be cited when used. Example “Appeltans W, Bouchet P, Boxshall GA, Fauchald K, Gordon DP, Hoeksema BW, Poore GCB, van Soest RWM,

⁹ http://links.gbif.org/gbif_metadata_profile_guide_en_v1

Stöhr S, Walter TC, Costello MJ. (eds) (2011). World Register of Marine Species.

Accessed at <http://www.marinespecies.org> on 2011-02-22."

- **Bibliography** - A complete bibliography of sources can be described and included in the metadata document.

Data-level Citation and Attribution

Attribution and citation information recorded in the metadata document is common to all data records in a dataset. In some cases, additional granularity is needed even down to individual records. In these cases, there are record-level terms that are recommended for use in specifying citation and attribution information.

- *dwc:nameAccordingTo* : This term can be used to identify the individual or citation that serves as the authoritative taxonomic reference for the record. (Example "Erpenbeck, D.; Van Soest, R.W.M. 2002. Family Halichondriidae Gray, 1867. Pp. 787-816. In Hooper, J. N. A. & Van Soest, R. W. M. (ed.) Systema Porifera. A guide to the classification of sponges.")
- *dwc:nameAccordingToID*: A unique identifier that returns the nameAccordingTo reference as described above. This could be a URL for example.
- *dwc:datasetName*: If the record is derived from an external dataset this dataset can be cited as a text string. (Example, "World Register of Marine Species, cited on 12 April 2011")
- *dwc:datasetID* - An identifier that refers to a dataset, preferably resolvable.
- *dc:source* - Link to the source web page

Use Case #1 - Checklists composed of multiple contributing datasets (e.g., Catalogue of Life, PESI, WoRMS)

A taxonomic dataset may be a composite of multiple contributing sources, each of which needs to be acknowledged in addition to the collective resource itself. There are many examples of this. Perhaps the largest such collective effort is the Catalogue of Life Annual Checklist which aims to provide a complete listing of all the worlds living species. The checklist itself is composed of individual datasets that represent major taxonomic groups. Each of these resources, in turn, may be composed of contributions from a sub-network of specialists.

Other examples include the Pan-European Species list, which is composed of a number of contributing datasets that include Fauna Europaea, the European Register of Marine

Species, Euro+Med PlantBase and others. The World Register of Marine Species represents another such network.

The recommended practice for effectively documenting the provenance of these sorts of resources can be summarized as follows.

1. A single metadata document is created to represent the collective resource itself, (e.g., the Catalogue of Life, the The World Register of Marine Species, etc.) This metadata document provides the proper citation, agents, rights, and other elements identified above. This document filename is referenced the Darwin Core Archive descriptor file, meta.xml. This links the document to the entire DwC-A dataset. Recommended best practice is that this file uses the GBIF metadata profile and be named EML.xml. In this case, the metadata descriptor XML would look like this:
 - a. `<archive xmlns="http://rs.tdwg.org/dwc/text/" metadata="eml.xml">`
2. Additional metadata documents can be created for each of the component datasets and included in the archive. This allows each sub-component dataset to be documented as completely as the “parent” dataset with its own recommended citation, contributing individuals etc. As these datasets do not document the entire collection, they are not referenced in the meta.xml descriptor file. Instead they are referenced from individual data records via the *dwc:datasetID* term. If the metadata documents are included in the archive itself, the *dwc:datasetID* equals the filename of the document. Alternatively, it could refer to a URL or some other unique and resolvable identifier for the information. A less recommended but alternative approach would be adding a URL to a simple web page that describes the dataset as opposed to a structured metadata document.
3. To cite individuals at the record level, providing a 3rd level of citation, it is recommended to use the *dwc:nameAccordingTo* field. Additional record-level terms are provided above.

Use Case #2 - Checklists derived from one or more authority sources

A species checklist in this use case is compiled for a specific purpose but derives it's basic taxonomic structure from one or more external taxonomic checklists that serve as *authority files*. The new compilation may include additional annotations to the basic source record that apply to the new lists focus. An example might be a European national species checklist derived from a database such as Fauna Europaea or the Catalogue of Life, which, in principle, provide the complete listing for a country as a subset of their own coverage. A national list may then add additional regional details such as a national

threat status or some other property of interest, resulting in a new, derived dataset. In this case, it is important to be able to provide record-level attribution and linkages to the source dataset. The recommended means to do this are as follows.

1. A single metadata document is created to represent the new, derived resource itself (e.g., National Checklist of the Netherlands). Datasets that are referenced can be cited in this metadata document.
 - A. Fully described as organisations with a role of Contributor and links to the source web site.
 - B. Cited in the bibliographic section with the citation represented as recommended by the referenced dataset.
2. In the datafiles, additional attribution and linkages can be made at the record-level. This includes:
 - a. Referencing the dataset by name in *dwc:datasetName*
 - b. Referencing the dataset by ID (such as URL) in *dwc:datasetID* and linking to the dataset home page
 - c. Providing a link to a corresponding species page on the referenced dataset web site using *dc:source*
 - i. If *dc:source* is reserved for pointing to URL for the derived database, a link to the source database can still be added using the Alternative Identifiers extension.
 - d. If the source dataset provides globally unique identifiers for the taxa referenced in the list, they can be used as the *taxonID* in the derived dataset. This ensures an explicit link to the source taxon and is highly recommended if available.
 - e. Use the *dwc:nameAccordingTo* or *dwc:nameAccordingToID* to refer to the taxon definition in the corresponding source record as a citation or a URL.

Authenticated File Access via httpd

GBIF advocates free and open access to biodiversity data and extends this advocacy to include taxonomic information. Some taxonomic data publishers, notwithstanding, may balk at providing data via a completely open and public URL, without an initial consultation with a data user, even if the data are free and open. For such a case, it is possible to publish a Darwin Core Archive such that it is only accessible via authentication

with a username and password. The URL itself may be published freely as authentication is required to access the file. Data publishers may use the web server logging function to track access via specific users. It is worth repeating that this is not the preferred method for GBIF. GBIF's position is that if there is a demand for taxonomic resources, a consistent and user-friendly citation and attribution process, such as has been defined here, is preferred.

The Core Taxon Class

Use the core class to provide the fundamental information for a species checklist including classification, synonymy and other key elements. Each row in the list can represent an accepted taxon or a synonym. The terms in this class support different methods for representing classification information.

RowType: <http://rs.tdwg.org/dwc/terms/Taxon>

Repository: http://rs.gbif.org/core/dwc_taxon.xml

taxonID	<p>A unique identifier for the set of nomenclatural and taxonomic information (data associated with the Taxon class). A taxonID refers to a name being used somewhere, therefore not precluding whether it is a taxon (concept) or name alone that is being described. This also includes both accepted taxon references and references to synonyms. If possible a globally unique ID is preferred, but local IDs are valid as long as they are unique within the dataset. It is possible for two different taxonID rows to refer to the same scientificName if they refer to different taxon concept circumscriptions referenced in nameAccordingTo.</p> <p>Example: "8fa58e08-08de-4ac1-b69c-1235340b7001; 32567; http://species.gbif.org/abies_alba_1753; urn:lsid:gbif.org:usages:32567"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/taxonID Data Type: string Required: false</p>
acceptedNameUsageID	<p>The taxonID of the taxon considered to be the valid (zoological) or accepted (botanical) name for this nameUsage. If acceptedNameUsage is given, the scientificName of the nameUsage with taxonID=acceptedNameUsageID needs to be the same as acceptedNameUsage. <i>This is the preferred method over using a verbatim acceptedNameUsage.</i></p> <p>Example: 8fa58e08-08de-4ac1-b69c-1235340b7001</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/acceptedNameUsageID Data Type: string Required: false</p>
acceptedNameUsage	<p>The scientificName of the taxon considered to be the valid (zoological) or accepted (botanical) name for this nameUsage.</p> <p>Example: "Tamias minimus" valid name for "Eutamias minimus"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/acceptedNameUsage Data Type: string Required: false</p>
parentNameUsageID	<p>The taxonID of the direct, most proximate higher-rank parent taxon (in a classification). This is the preferred method for referencing the parent taxon (preferred over repeating the parent taxon verbatim using parentNameUsage below)</p> <p>Example: 8fa58e08-08de-4ac1-b69c-1235340b7001</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/parentNameUsageID Data Type: string Required: false</p>
parentNameUsage	<p>The scientificName representing the direct, most proximate higher-rank parent taxon (in a taxonomic classification) of this taxon's scientificName. This name must match the name used in the row representing the parent taxon.</p> <p>Example: "Rubiaceae", "Gruiformes", "Testudinae"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/parentNameUsage Data Type: string Required: false</p>

originalNameUsageID	A unique identifier for the nameUsage instance in which the name was originally established, under the rules of the associated nomenclaturalCode (i.e., within the namePublishedIn reference). The basionym (botany) or basonym (bacteriology) of the scientificName or the senior/earlier homonym for replaced names. If provided the nameAccordingTo value returned should match the namePublishedIn value for this record. Example: http://species.gbif.org/abies_alba_1753 Qualified Name: http://rs.tdwg.org/dwc/terms/originalNameUsageID Data Type: string Required: false
originalNameUsage	The equivalent of the scientificName as it originally appeared when the name was first established under the rules of the associated nomenclaturalCode (i.e., within the namePublishedIn reference). The basionym (botany) or basonym (bacteriology) of the scientificName or the senior/earlier homonym for replaced names. Example: "Pinus abies", "Gasterosteus saltatrix Linnaeus 1768" Qualified Name: http://rs.tdwg.org/dwc/terms/originalNameUsage Data Type: string Required: false
nameAccordingTo	To use a name precisely an indication of which concept of that name one refers to is needed. Traditionally the Latin sensu or sec. (for secundum - according to) have been used. For taxa that result from identifications a reference to the keys used, monographs, online source or experts should be given. If no accordingTo is explicitly given the "nominal concept" as defined by TCS is assumed. Could be a publication (identification key), institution or team of individuals. Example: "Werner Greuter 2008; Lilljeborg 1861, Upsala Univ. Arsskrift, Math. Naturvet., pp. 4, 5", "McCrane, J. R., D. B. Wake, and L. D. Wilson. 1996. The taxonomic status of Bolitoglossa schmidti, with comments on the biology of the Mesoamerican salamander Bolitoglossa dofleini (Caudata: Plethodontidae). Carib. J. Sci. 32:395-398." Qualified Name: http://rs.tdwg.org/dwc/terms/nameAccordingTo Data Type: string Required: false
nameAccordingToID	A unique identifier that returns the details of a nameAccordingTo reference. Example: "doi:10.1016/S0269-915X(97)80026-2" Qualified Name: http://rs.tdwg.org/dwc/terms/nameAccordingToID Data Type: string Required: false
namePublishedInID	A preferably resolvable, globally unique identifier that refers to namePublishedIn. http://hdl.handle.net/10199/7 Qualified Name: http://rs.tdwg.org/dwc/terms/namePublishedInID Data Type: string Required: false
namePublishedIn	Reference to a publication representing the original publication of the name. Example: "Forel, Auguste, Diagnoses provisoires de quelques espèces nouvelles de fourmis de Madagascar, récoltées par M. Grandidier., Annales de la Societe Entomologique de de Belgique, Comptes-rendus des Seances 30, 1886" Qualified Name: http://rs.tdwg.org/dwc/terms/namePublishedIn Data Type: string Required: false
taxonConceptID	Qualified Name: http://rs.tdwg.org/dwc/terms/taxonConceptID Data Type: string Required: false
scientificNameID	Exclusively used to reference an external and resolvable identifier that returns nomenclatural (not taxonomic) details of a name. Use taxonID to refer to taxa.

	<p>Example: "urn:lsid:ipni.org:names:37829-1:1.3"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/scientificNameID</p> <p>Data Type: string</p> <p>Required: false</p>
scientificName	<p>The taxon name (with date and authorship information if applicable). When forming part of an Identification, this should be the name in lowest level taxonomic rank that can be determined. This term should not contain identification qualifications, which should instead be supplied in the IdentificationQualifier term.</p> <p>Example: "Coleoptera" (order), "Vespertilionidae" (family), "Manis" (genus), "Ctenomys sociabilis" (genus + specificEpithet), "Ambystoma tigrinum diaboli" (genus + specificEpithet + infraspecificEpithet), "Quercus agrifolia var. oxyadenia (Torr.)"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/scientificName</p> <p>Data Type: string</p> <p>Required: false</p>
scientificNameAuthorship	<p>The authorship information for the scientificName formatted according to the conventions of the applicable nomenclaturalCode.</p> <p>Example: "(Torr.) J.T. Howell", "(Martinovsk) Tzvelev", "(Linnaeus 1768)"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/scientificNameAuthorship</p> <p>Data Type: string</p> <p>Required: false</p>
higherClassification	<p>A list (concatenated and separated) of taxa names terminating at the rank immediately superior to the taxon referenced in the taxon record. Recommended best practice is to order the list starting with the highest rank and separating the names for each rank with a semi-colon (";").</p> <p>Example: Animalia;Chordata;Vertebrata;Mammalia;Theria;Eutheria;Rodentia;Hystricognatha;Hystricognathi;Ctenomyidae;Ctenomyini;Ctenomys</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/higherClassification</p> <p>Data Type: string</p> <p>Required: false</p>
kingdom	<p>The full scientific name of the kingdom in which the taxon is classified.</p> <p>Example: "Animalia", "Plantae"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/kingdom</p> <p>Data Type: string</p> <p>Required: false</p>
phylum	<p>The full scientific name of the phylum in which the taxon is classified.</p> <p>Example: "Chordata" (phylum), "Bryophyta" (division)</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/phylum</p> <p>Data Type: string</p> <p>Required: false</p>
class	<p>The full scientific name of the class in which the taxon is classified.</p> <p>Example: "Mammalia", "Hepaticopsida"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/class</p> <p>Data Type: string</p> <p>Required: false</p>
order	<p>The full scientific name of the order in which the taxon is classified.</p> <p>Example: "Carnivora", "Monocleales"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/order</p> <p>Data Type: string</p> <p>Required: false</p>

family	<p>The full scientific name of the family in which the taxon is classified.</p> <p>Example: "Felidae", "Monocleaceae"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/family Data Type: string Required: false</p>
genus	<p>The full scientific name of the genus in which the taxon is classified.</p> <p>Example: "Puma", "Monoclea"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/genus Data Type: string Required: false</p>
subgenus	<p>The full scientific name of the subgenus in which the taxon is classified. Values should include the genus to avoid homonym confusion.</p> <p>Example: Puma (Puma); Loligo (Amerigo); Hieracium subgen. Pilosella</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/subgenus Data Type: string Required: false</p>
specificEpithet	<p>The name of the species epithet of the scientificName.</p> <p>Example: "concolor", "gottschei"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/specificEpithet Data Type: string Required: false</p>
infraspecificEpithet	<p>The name of the lowest or terminal infraspecific epithet of the scientificName, excluding any rank marker.</p> <p>Example: "concolor", "oxyadenia", "sayi"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/infraspecificEpithet Data Type: string Required: false</p>
taxonRank	<p>The taxonomic rank of the most specific name in the scientificName. Recommended vocabulary: http://rs.gbif.org/vocabulary/gbif/rank.xml</p> <p>Example: "subspecies", "varietas", "forma", "species", "genus"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/taxonRank Data Type: string Required: false</p>
verbatimTaxonRank	<p>The taxonomic rank of the most specific name in the scientificName as it appears in the original record or the rank designator within the verbatim original name itself. May include abbreviations for example.</p> <p>Example: "Agamospecies", "sub-lesus", "prole", "apomict", "nothogrex", "sp.", "subsp.", "var."</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/verbatimTaxonRank Data Type: string Required: false</p>
vernacularName	<p>A common or vernacular name.</p> <p>Example: "Andean Condor", "Condor Andino", "American Eagle", "Gänsegeier"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/vernacularName Data Type: string Required: false</p>
nomenclaturalCode	<p>The nomenclatural code under which the scientificName is constructed.</p> <p>Example: ICBN; ICZN</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/nomenclaturalCode Data Type: http://rs.gbif.org/vocabulary/gbif/nomenclatural_code Required: false</p>

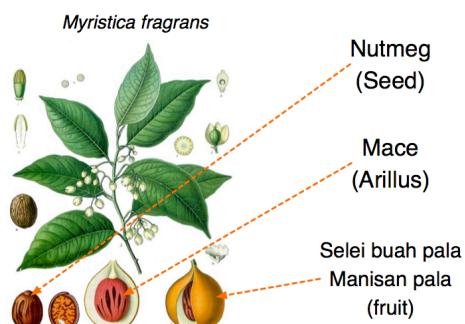
taxonomicStatus	<p>The status of the use of the scientificName as a label for a taxon. Requires taxonomic opinion to define the scope of a taxon. Rules of priority then are used to define the taxonomic status of the nomenclature contained in that scope, combined with the experts opinion. It must be linked to a specific taxonomic reference that defines the concept. Recommended vocabulary:</p> <p>http://rs.gbif.org/vocabulary/gbif/taxonomic_status.xml</p> <p>Example: "invalid", "misapplied", "homotypic synonym", "accepted"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/taxonomicStatus Data Type: string Required: false</p>
nomenclaturalStatus	<p>The status related to the original publication of the name and its conformance to the relevant rules of nomenclature. It is based essentially on an algorithm according to the business rules of the code. It requires no taxonomic opinion. Recommended vocabulary:</p> <p>http://rs.gbif.org/vocabulary/gbif/nomenclatural_status.xml</p> <p>Example: "nom. ambig.", "nom. illeg.", "nom. subnud."</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/nomenclaturalStatus Data Type: string Required: false</p>
taxonRemarks	<p>Comments or notes about the taxon or name.</p> <p>Example: Type consists of a skull and skeletal fragments.</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/taxonRemarks Data Type: string Required: false</p>
modified	<p>Date when the record was last updated</p> <p>Example: 2009-08-21</p> <p>Qualified Name: http://purl.org/dc/terms/modified Data Type: date Required: false</p>
language	<p>A language of the resource. Recommended best practice is to use a controlled vocabulary such as ISO 693</p> <p>Example: "eng"</p> <p>Qualified Name: http://purl.org/dc/terms/language Data Type: string Required: false</p>
rights	<p>Information about rights held in and over the resource. Typically, rights information includes a statement about various property rights associated with the resource, including intellectual property rights.</p> <p>Example: "Content licensed under Creative Commons Attribution 3.0 United States License", "CC BY-SA"</p> <p>Qualified Name: http://purl.org/dc/terms/rights Data Type: string Required: false</p>
rightsHolder	<p>A person or organization owning or managing rights over the resource.</p> <p>Qualified Name: http://purl.org/dc/terms/rightsHolder Data Type: string Required: false</p>
accessRights	<p>Information about who can access the resource or an indication of its security status. Access Rights may include information regarding access or restrictions based on privacy, security, or other policies.</p> <p>Example: "not-for-profit use only"</p> <p>Qualified Name: http://purl.org/dc/terms/accessRights Data Type: string</p>

	Required: false
bibliographicCitation	<p>Citation information specified by the data publisher. Citation information is inherited downward by all child taxa if no other citation is included. Citation information is NOT accumulated upward. For example, one citation may be linked to a Mammalia entry and generally applies to all mammal species but a different citation for a child taxon, Primates, applies to all child primate taxa.</p> <p>Example: “van Soest, R. (2009). <i>Leucandra fistulosa</i> (Johnston, 1842). In: Van Soest, R.W.M., Boury-Esnault, N., Hooper, J.N.A., Rützler, K., de Voogd, N.J., Alvarez, B., Hajdu, E., Pisera, A.B., Vacelet, J., Manconi, R., Schoenberg, C., Janussen, D., Tabachnick, K.R., Klautau, M. (Eds) (2009). World Porifera database”</p> <p>Qualified Name: http://purl.org/dc/terms/bibliographicCitation Data Type: string Required: false</p>
informationWithheld	<p>Additional remarks as to information not published, but available</p> <p>Example: “hybrid parents information available”</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/informationWithheld Data Type: string Required: false</p>
datasetID	<p>An identifier for a (sub) dataset. Ideally globally unique, but any id allowed</p> <p>Example: “13”</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/datasetID Data Type: string Required: false</p>
datasetName	<p>The title of the (sub)dataset optionally also referenced via datasetID</p> <p>Example: “World Register of Marine Species” “Fauna Europaea”</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/datasetName Data Type: string Required: false</p>
source	<p>A URI link or reference to the source of this record. A link to a webpage or RESTful webservice is recommended. URI is mandatory format.</p> <p>Example: “http://www.catalogueoflife.org/annual-checklist/show_species_details.php?record_id=6197868”</p> <p>Qualified Name: http://purl.org/dc/terms/source Data Type: uri Required: false</p>

Sharing Vernacular Names

The GNA Profile supports the sharing of vernacular name data associated with taxa in taxonomic checklists. Vernacular names are shared as a separate, related file using the [Vernacular Names extension](#)¹⁰. The extension supports a rich set of properties for describing vernacular name usages that include regional and morphological qualifiers. The complete listing of extension terms and recommended vocabularies can be found in the GBIF Resource Repository referenced in the previous footnote or in the [GNA Profile Reference Guide](#)¹¹.

Vernacular names are referenced via an extension, therefore they must be linked to a named taxon in the parent core data file. It is further recommended that a vernacular names record provide a language reference that identifies the language represented by the vernacular name use. The best practice is to use the ISO 693 language codes for sharing language information. The complete set of language codes can be found on the [GBIF vocabulary server](#)¹². Vernacular names may also have distinct regional uses and this can be specified through a dwc:locality element or, at a less precise level, using a dwc:country term. It is recommended that country names utilise the ISO 6133 country codes, which are also available on the [GBIF vocabulary server](#)¹³.



Vernacular Names Extension

This extension provides the means to share information related to common (vernacular) names linked to taxa in the core data file. Multiple vernacular names can be linked to the same taxon via the taxonID.

RowType: <http://rs.gbif.org/terms/1.0/VernacularName>
 Repository: <http://rs.gbif.org/extension/gbif/1.0/vernacular.xml>

taxonID	The identifier used in the core data file representing the taxon for which the current record refers. This identifier provides the link
---------	---

¹⁰ Vernacular Names Extension - <http://rs.gbif.org/extension/gbif/1.0/vernacularname.xml>

¹¹ http://links.gbif.org/gbif_gna_profile_reference_guide

¹² <http://vocabularies.gbif.org/vocabularies/lang>

¹³ <http://vocabularies.gbif.org/vocabularies/country>

	between the core data record and the extension record. Qualified Name: http://rs.tdwg.org/dwc/terms/taxonID Data Type: string Required: true
vernacularName	A common or vernacular name. Example: "Andean Condor", "Condor Andino", "American Eagle", "G/Snsegeier" Qualified Name: http://rs.tdwg.org/dwc/terms/vernacularName Data Type: string Required: true
source	Bibliographic citation referencing a source where the vernacular name refers to the cited species. Example: "Peterson Field Guide to the Eastern Seashore, Houghton Mifflin Co, 1961, p131" Qualified Name: http://purl.org/dc/terms/source Data Type: string Required: false
language	ISO 639-1 language code used for the vernacular name value. Example: "ES" Qualified Name: http://purl.org/dc/terms/language Data Type: http://iso.org/639-1 Required: false
temporal	temporal context when name is/was used Example: 19th Century; 1950 Qualified Name: http://purl.org/dc/terms/temporal Data Type: string Required: false
locationID	An identifier for the set of location information (data associated with dcterms:Location). May be a global unique identifier or an identifier specific to the data set. Qualified Name: http://rs.tdwg.org/dwc/terms/locationID Data Type: string Required: false
locality	The specific description of the area from which the vernacular name usage originates. Vernacular names may have very specific regional contexts. A name used for a species in one area may refer to a different species in another. Example: "Southeastern coastal New England from Buzzards Bay through Rhode Island" Qualified Name: http://rs.tdwg.org/dwc/terms/locality Data Type: string Required: false
countryCode	The standard code for the country in which the vernacular name is used. Recommended best practice is to use the ISO 3166-1-alpha-2 country codes available as a vocabulary at http://rs.gbif.org/vocabulary/iso/3166-1_alpha2.xml . For multiple countries separate values with a comma "," Example: "AR" for Argentina, "SV" for El Salvador. "AR,CR,SV" for Argentina, Costa Rica, and El Salvador combined. Qualified Name: http://rs.tdwg.org/dwc/terms/countryCode Data Type: string Required: false
sex	The sex (gender) of the taxon for which the vernacular name applies

	<p>when the vernacular name is limited to a specific gender of a species. If not limited sex should be empty. For example the vernacular name "Buck" applies to the "Male" gender of the species, <i>Odocoileus virginianus</i>.</p> <p>Example: "male"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/sex Data Type: http://rs.gbif.org/vocabulary/gbif/sex Required: false</p>
lifeStage	<p>The age class or life stage of the species for which the vernacular name applies. Best practice is to utilise a controlled list of terms for this value.</p> <p>Example: 'juvenile' is the life stage of the fish <i>Pomatomus saltatrix</i> for which the name "snapper blue" refers.'</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/lifeStage Data Type: http://rs.gbif.org/vocabulary/gbif/life_stage Required: false</p>
isPlural	<p>This value is true if the vernacular name it qualifies refers to a plural form of the name.</p> <p>Example: The term "Schoolies" is the plural form of a name used along the coastal Northeastern U.S. for groups of juvenile fish of the species, <i>Morone saxatilis</i>.</p> <p>Qualified Name: http://rs.gbif.org/terms/1.0/isPlural Data Type: http://rs.gbif.org/vocab/boolean/ Required: false</p>
isPreferredName	<p>This term is true if the source citing the use of this vernacular name indicates the usage has some preference or specific standing over other possible vernacular names used for the species.</p> <p>Some organisations have attempted to assign specific and unique vernacular names for particular taxon groups in a systematic attempt to bring order and consistency to the use of these names. For example, the American Ornithological Union assigns the name "Pearl Kite" for the taxon, <i>Gampsonyx swainsonii</i>. The value of isPreferredName for this record would be true.</p> <p>Example: "True" "False"</p> <p>Qualified Name: http://rs.gbif.org/terms/1.0/isPreferredName Data Type: http://rs.gbif.org/vocab/boolean/ Required: false</p>
organismPart	<p>The part of the organism to which the vernacular name refers. Best practice is to utilise a controlled vocabulary for this term although it is likely that multiple controlled lists for different organism groups may be the best implementation for this term.</p> <p>The spice "Mace", is derived from the "aril" of the plant <i>Myristica fragrans</i> while the spice "nutmeg" is derived from the "seed." "Seed" and "Aril" represent two different values for organismPart.</p> <p>Qualified Name: http://rs.gbif.org/terms/1.0/organismPart Data Type: string Required: false</p>
taxonRemarks	<p>A description of any context that qualify the specific usage of the vernacular name.</p> <p>Example: "This name applies only when the fruit has been blessed by the tribal shaman"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/taxonRemarks Data Type: string Required: false</p>

Sharing Species Descriptions

The GNA Profile supports the sharing of descriptive information related to a taxon via the [Taxon Description extension](#)¹⁴. Descriptive data can be assigned to distinct description types and, as the data is published in an extension, multiple descriptive records may be linked to a single taxon, supporting a relatively rich set of data per taxon. It is recommended that a [description type vocabulary](#)¹⁵ be used to describe the descriptive information.

Multi-line descriptions

Descriptive information should be limited to single paragraph text blocks. Multiple paragraphs containing line breaks should be avoided or carefully managed in order to maintain the integrity of the resultant text file output as the Darwin Core Archive. Multi-line data fields served as text files require the record delimiters, which are usually line break characters, to be distinct from the line breaks used within a multi-line field. The best method for supporting multiple lines in a single field is to replace breaking characters with a non-breaking character or character set that a user can replace with proper breaks when the data is parsed and used. One option is to use the HTML break tag “
.”

Species Description Extension

Use this extension to provide descriptive text for a taxon. This is typically in the form of a single paragraph per record as would be normally stored in a database. Descriptions can be qualified by a type to indicate, for example that the description is related to, for example, the morphology conservation, reproduction, etc.

RowType: <http://rs.gbif.org/terms/1.0/Description>

Repository: <http://rs.gbif.org/extension/gbif/1.0/description.xml>

taxonID	<p>The identifier used in the core data file representing the taxon for which the current record refers. This identifier provides the link between the core data record and the extension record.</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/taxonID Data Type: string Required: true</p>
description	<p>Any descriptive free text matching the category given as dc:type</p> <p>Qualified Name: http://purl.org/dc/terms/description Data Type: string</p>

¹⁴ <http://rs.gbif.org/extension/gbif/1.0/description.xml>

¹⁵ http://rs.gbif.org/vocabulary/gbif/description_type.xml

	Required: true
type	<p>The kind of description given. See the Description Type vocabulary at http://rs.gbif.org/vocabulary/gbif/description_type.xml for a recommended list of description types.</p> <p>Example: “morphology, distribution”</p> <p>Qualified Name: http://purl.org/dc/terms/type Data Type: http://rs.gbif.org/vocabulary/gbif/descriptionType Required: true</p>
source	<p>Source reference of this description, a URL or full publication citation</p> <p>Qualified Name: http://purl.org/dc/terms/source Data Type: string Required: false</p>
language	<p>ISO 639-1 language code used for the vernacular name value.</p> <p>Example: “ES”</p> <p>Qualified Name: http://purl.org/dc/terms/language Data Type: string Required: false</p>
creator	<p>The author(s) of the textual information provided for a description</p> <p>Example: “Hershkovitz, P.”</p> <p>Qualified Name: http://purl.org/dc/terms/creator Data Type: string Required: false</p>
contributor	<p>An entity responsible for making contributions to the textual information provided for a description</p> <p>Qualified Name: http://purl.org/dc/terms/contributor Data Type: string Required: false</p>
audience	<p>A class or description for whom the dwc:description is intended or useful</p> <p>Example: “experts”, “general public”, “children”</p> <p>Qualified Name: http://purl.org/dc/terms/audience Data Type: string Required: false</p>
license	<p>Official permission to do something with the resource. Please use Creative Commons URLs if you can.</p> <p>Example: CC-BY</p> <p>Qualified Name: http://purl.org/dc/terms/license Data Type: string Required: false</p>
rightsHolder	<p>A person or organization owning or managing rights over the resource.</p> <p>Qualified Name: http://purl.org/dc/terms/rightsHolder Data Type: string Required: false</p>

Sharing Species Distributions

The GNA Profile supports the sharing of distribution data via the [Species Distribution extension](#)¹⁶. This enables multiple distribution records to be published per taxon. The distribution extension is not only used to designative national or regional distribution descriptions, it also supports the qualification of the referenced distribution in regard to the threat status of the taxon, whether it is introduced, native, etc., and other properties that might be tied to a specific defined area.

The recommended best practice for specifying a distinct area is via a resolve-able or well-known area identifier published via the dwc:localityID element.

Species Distribution Extension

Use this extension to share information on one or more distribution references for a taxon. One ore more locality records may be linked to the same taxon. For example multiple localities, regions, or countries may be listed. Use this extension to describe the threat status for a taxon, seasonal distribution changes, and other properties linked to a taxon in a particular region.

RowType: <http://rs.gbif.org/terms/1.0/Distribution>

Repository: <http://rs.gbif.org/extension/gbif/1.0/distribution.xml>

taxonID	<p>The identifier used in the core data file representing the taxon for which the current record refers. This identifier provides the link between the core data record and the extension record.</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/taxonID Data Type: string Required: true</p>
locationID	<p>A code for the named area this distributon record is about. See http://en.wikipedia.org/wiki/ISO_3166-2 for state codes within a country, otherwise use a prefix for each code to indicate the source of the code</p> <p>Example: “TDWG:AGS-TF; ISO3166:AR; WOEID:564721”</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/locationID Data Type: string Required: false</p>
locality	<p>The verbatim name of the area this distributon record is about.</p> <p>Example: “Patagonia”</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/locality Data Type: string Required: false</p>

¹⁶ <http://rs.gbif.org/extension/gbif/1.0/distribution.xml>

countryCode	ISO3166 alpha 2 (3 is permissible) country codes the area belongs to or as an alternative for a locationID if the area is a country. For multiple countries separate values with a comma "," Example: "AR" "US;CA;MX" Qualified Name: http://rs.tdwg.org/dwc/terms/countryCode Data Type: string Required: false
lifeStage	The distribution information pertains solely to a specific life stage of the taxon. See the Life Stage vocabulary at http://rs.gbif.org/vocabulary/gbif/life_stage.xml for recommended values. Example: "adult" Qualified Name: http://rs.tdwg.org/dwc/terms/lifeStage Data Type: http://rs.gbif.org/vocabulary/gbif/life_stage Required: false
occurrenceStatus	Term describing the status of the organism in the given area based on how frequent the species occurs. See the Occurrence status vocabulary at http://rs.gbif.org/vocabulary/gbif/occurrence_status.xml for recommended values. Example: "Absent" "present" Qualified Name: http://rs.tdwg.org/dwc/terms/occurrenceStatus Data Type: http://rs.gbif.org/vocabulary/gbif/occurrence_status/ Required: false
threatStatus	Threat status of a species as defined by IUCN: http://www.iucnredlist.org/static/categories_criteria_3_1#categories Example: "EX" "EW" "CR" Qualified Name: http://rs.gbif.org/terms/1.0/threatStatus Data Type: http://rs.gbif.org/vocabulary/iucn/threat_status/ Required: false
establishmentMeans	Term describing whether the organism occurs natively, is introduced or cultivated. Example: "introduced" Qualified Name: http://rs.tdwg.org/dwc/terms/establishmentMeans Data Type: http://rs.gbif.org/vocabulary/gbif/establishment_means/ Required: false
appendixCITES	The CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora) Appendix number the taxa is listed. It is possible to have different appendix numbers for different areas, but "global" as an area is also valid if its the same worldwide Example: "II" Qualified Name: http://rs.gbif.org/terms/1.0/appendixCITES Data Type: http://rs.gbif.org/vocab/cites/ Required: false
eventDate	Relevant temporal context for this entire distribution record including all properties preferably given as a year range or single year on which the distribution record is valid. For the same area and taxon there could therefore be several records with different temporal context, e.g. in 5 year intervals for invasive species. Example: "1930"; "1939-1945" Qualified Name: http://rs.tdwg.org/dwc/terms/eventDate Data Type: string Required: false
startDayOfYear	Seasonal temporal subcontext within the eventDate context. Useful for

	<p>migratory species. The earliest ordinal day of the year on which the distribution record is valid. Numbering starts with 1 for January 1 and ends with 365 for December 31.</p> <p>Example: "90"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/startDayOfYear Data Type: string Required: false</p>
endDayOfYear	<p>Seasonal temporal subcontext within the eventDate context. The latest ordinal day of the year on which the distribution record is valid</p> <p>Example: "120"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/endDayOfYear Data Type: string Required: false</p>
source	<p>Source reference for this distribution record. Can be proper publication citation, a webpage URL, etc.</p> <p>Example: "Euro+Med Plantbase - the information resource for Euro-Mediterranean plant diversity (2006-). Published on the Internet http://ww2.bgbm.org/EuroPlusMed/ July, 2009"</p> <p>Qualified Name: http://purl.org/dc/terms/source Data Type: string Required: false</p>
occurrenceRemarks	<p>Comments or notes about the distribution</p> <p>Example: "Excluded because of misidentification"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/occurrenceRemarks Data Type: string Required: false</p>

If the dwc:country element is used, it is recommended that the ISO 3166 country codes, available on the [GBIF vocabulary server](http://vocabularies.gbif.org/vocabularies/country)¹⁷, be used.

Sharing References

The GNA Profile supports the sharing of bibliographic citations through the [References extension](#)¹⁸. The References extension is recommended and designed for use in the sharing of synonymy information in monographs and annotated checklists. It supports the sharing of a parsed citation and therefore provides a more granular citation format than some of the citation-storing data elements in the core data file, such as dwc:namePublishedIn. This extension supports the taxonomic and nomenclatural qualification of a reference via the dc:type property, which, when used with [the Reference Type vocabulary](#)¹⁹, can be used to distinguish a set of references related to a taxon.

References Extension

Use this extension to describe one or more bibliographic references related to a taxon in the core data file. Use the type field to qualify the references. This extension supports the sharing of referenced synonymous checklists.

RowType: <http://rs.gbif.org/terms/1.0/Reference>

Repository: <http://rs.gbif.org/extension/gbif/1.0/references.xml>

taxonID	<p>The identifier used in the core data file representing the taxon for which the current record refers. This identifier provides the link between the core data record and the extension record.</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/taxonID Data Type: string Required: true</p>
identifier	<p>DOI, ISBN, URI, etc referring to the reference. This can be repeated in multiple rows to include multiple identifiers, e.g. a DOI and a URL pointing to a pdf of the article.</p> <p>Example: doi:10.1038/ng0609-637; http://www.nature.com/ng/journal/v41/n6/pdf/ng0609-637.pdf</p> <p>Qualified Name: http://purl.org/dc/terms/identifier Data Type: string Required: false</p>
bibliographicCitation	<p>A text string referring to an un-parsed bibliographic citation.</p> <p>Example: "Hartge, P., Genetics of reproductive lifespan. Nature</p>

¹⁷ <http://vocabularies.gbif.org/vocabularies/country>

¹⁸ <http://rs.gbif.org/extension/gbif/1.0/references.xml>

¹⁹ http://rs.gbif.org/vocabulary/gbif/reference_type.xml

	<p>Genetics 41, 637 - 638 (2009)"</p> <p>Qualified Name: http://purl.org/dc/terms/bibliographicCitation Data Type: string Required: false</p>
title	<p>Title of book or article</p> <p>Example: "Genetics of reproductive lifespan", "Field Guide to Moths of Eastern North America"</p> <p>Qualified Name: http://purl.org/dc/terms/title Data Type: string Required: false</p>
creator	<p>The author or authors of the referenced work</p> <p>Example: "Patricia Hartge"</p> <p>Qualified Name: http://purl.org/dc/terms/creator Data Type: string Required: false</p>
date	<p>Date of publication, recommended ISO format YYYY or YYYY-MM-DD</p> <p>Example: "6/1/2009"; "2009"</p> <p>Qualified Name: http://purl.org/dc/terms/date Data Type: string Required: false</p>
source	<p>If the reference is part of a larger work, this can be cited here. In case of articles this is the journal, for parts of books the book itself</p> <p>Example: Nature Genetics 41, 635 (2009)</p> <p>Qualified Name: http://purl.org/dc/terms/source Data Type: string Required: false</p>
description	<p>Abstracts, remarks, notes</p> <p>Example: "Five genome-wide association studies of the timing of menarche and menopause have now taken us beyond the range of candidate gene and linkage studies. The list of new genetic associations identified for these two traits should shed light on the mechanisms of ovarian aging, as well as breast cancer and other diseases associated with reproductive lifespan."</p> <p>Qualified Name: http://purl.org/dc/terms/description Data Type: string Required: false</p>
subject	<p>Semicolon seperated list of keywords. Can include a resource qualifier that specifies the relation of this reference to the taxon, e.g namePublishedIn</p> <p>Example: genomics; epidemiology</p> <p>Qualified Name: http://purl.org/dc/terms/subject Data Type: string Required: false</p>
language	<p>ISO 639-1 language code indicating the source language of the referent publication</p> <p>Example: "en"</p> <p>Qualified Name: http://purl.org/dc/terms/language Data Type: http://iso.org/639-1 Required: false</p>
rights	<p>copyright information relating to the referenced publication</p> <p>Example: "Copyright © 2009 Wiley-Liss, Inc., A Wiley Company"</p> <p>Qualified Name: http://purl.org/dc/terms/rights</p>

	<p>Data Type: string Required: false</p>
taxonRemarks	<p>Annotation of taxon-specific information related to the referenced publication.</p> <p>Example: "transferred H. nigritarsus to Acanolonia"; "Type specimen is a skeleton"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/taxonRemarks Data Type: string Required: false</p>
type	<p>Used to assign a bibliographic reference to list of taxonomic or nomenclatural categories. Best practice is to use a controlled vocabulary. See an example below in data type.</p> <p>Example: "Original publication of new combination (comb nov.)"</p> <p>Qualified Name: http://purl.org/dc/terms/type Data Type: http://rs.gbif.org/vocabulary/gbif/referenceType Required: false</p>

Sharing Type information

The GNA Profile supports the sharing of information about types and specimens via the Types and Specimens extension²⁰. It supports the sharing of basic information about type specimens, type species and genera.

Types and Specimens Extension

Use this extension to share data relating to one or more specimens or type references linked to the core taxon

RowType: <http://rs.gbif.org/terms/1.0/Specimen>

Repository: http://rs.gbif.org/sandbox/extension/gbif_taxon_type_specimens.xml

taxonID	<p>The identifier used in the core data file representing the taxon for which the current record refers. This identifier provides the link between the core data record and the extension record.</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/taxonID Data Type: string Required: true</p>
bibliographicCitation	<p>A text string citating the described specimen. Often found in taxonomic treatments and frequently based on institution code and catalog number.</p> <p>Example: Iraq: Mosul: Jabal Khantur prope Sharanish N. Zakho, in fissures rupium calc., 1200 m, Rech. 12083 (W!).</p> <p>Qualified Name: http://purl.org/dc/terms/bibliographicCitation Data Type: string Required: false</p>
occurrenceID	<p>An identifier for the specimen, preferably a resolvable globally unique identifier.</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/occurrenceID Data Type: string Required: false</p>
institutionCode	<p>The name (or acronym) in use by the institution having custody of the object(s) or information referred to in the record.</p> <p>Examples: "MVZ", "FMNH", "AKN-CLO", "University of California Museum of Paleontology (UCMP)"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/institutionCode Data Type: string Required: false</p>
collectionCode	<p>The name, acronym, coden, or initialism identifying the collection or data set from which the record was derived.</p> <p>Examples: "Mammals", "Hildebrandt", "eBird"</p>

²⁰ <http://rs.gbif.org/extension/gbif/1.0/typesandspecimen.xml>

	<p>Qualified Name: http://rs.tdwg.org/dwc/terms/collectionCode Data Type: string Required: false</p>
catalogNumber	<p>An identifier (preferably unique) for the record within the data set or collection. Examples: "2008.1334", "145732a", "145732"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/catalogNumber Data Type: string Required: false</p>
typeStatus	<p>The type status of the specimen. Preferably taken from a vocabulary like http://rs.gbif.org/vocabulary/gbif/type_status.xml</p> <p>Example: "holotype" "syntype" "lectotype"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/typeStatus Data Type: string Required: false</p>
scientificName	<p>The scientific name as which this specimen has been identified in the collection/source. Not necessarily the same as taxonID</p> <p>Example: <i>Ctenomys sociabilis</i> "Roptrocerus typographi (Gyðrfi, 1952)"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/scientificName Data Type: string Required: false</p>
taxonRank	<p>The rank of the taxon bearing the scientific name</p> <p>Example: "subspecies", "varietas", "forma", "species", "genus"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/taxonRank Data Type: string Required: false</p>
identificationRemarks	<p>Information regarding the basis of the identification or designation (in the case of type species and type genera)</p> <p>Example: by monotypy</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/identificationRemarks Data Type: string Required: false</p>
locality	<p>The location where the specimen was collected. In case of type specimens the type locality.</p> <p>Example: Iraq: Mosul: Jabal Khantur prope Sharanish N. Zakho, in fissures rupium calc., 1200 m</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/locality Data Type: string Required: false</p>
sex	<p>The sex of the specimen being referenced.</p> <p>Example: male</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/sex Data Type: string Required: false</p>
recordedBy	<p>The primary collector or observer, especially one who applies a personal identifier (recordNumber), should be listed first.</p> <p>Example: KH Rechinger</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/recordedBy Data Type: string Required: false</p>
source	Source reference for this type record. Can be proper publication

	citation, a webpage URL, etc. Qualified Name: http://purl.org/dc/terms/source Data Type: string Required: false
verbatimEventDate	The date when the specimen was collected Example: "spring 1910", "Marzo 2002", "1999-03-XX", "17IV1934" Qualified Name: http://rs.tdwg.org/dwc/terms/verbatimEventDate Data Type: string Required: false
verbatimLabel	The full, verbatim text from the specimen label Qualified Name: http://rs.gbif.org/terms/1.0/verbatimLabel Data Type: string Required: false
verbatimLongitude	The geographic longitude Example: 121d 10' 34 W Qualified Name: http://rs.tdwg.org/dwc/terms/verbatimLongitude Data Type: string Required: false
verbatimLatitude	The geographic latitude Example: 41 05 54.03S Qualified Name: http://rs.tdwg.org/dwc/terms/verbatimLatitude Data Type: string Required: false

Sharing Links and Identifiers

The GNA profile supports the means to share and [describe multiple links to related external resources](#)²¹. It allows data publishers to embed links back to the source database or document via resolve-able identifiers. Multiple identifiers, perhaps linking to both a web page as well as a more machine-readable web service response, may be provided for a single taxon. It is recommended that a format be included for each record to enable a user to know how to interpret the response information if an identifier is resolve-able. This is usually done by including the *mime type* in this field. A complete list of mime types is [available](#)²².

Creating a dynamic link to a species page

Often, a link back to a source database follows a common format, differing only in the identifier number or taxon name used in the URL. This can result in a verbose and bloated extension file. The DarwinCore Archive format supports a more efficient way to define a URL template, which only needs to be defined once, and allows a variable to be embedded in the template eliminating the need for repetitively repeating a set of URLs for each taxon in the data file. This is done via the XML metafile component of a DarwinCore Archive. It does not use the References extension. This requires editing the XML metafile which requires some degree of familiarity with XML. GBIF [provides a complete guide the to Darwin Core metafile](#)²³.

The metafile supports the creation of variables in the metafile that may refer to a web page or web service call. This variable may be embedded in the URL and include a taxon identifier or the taxon name as one of the parameters in the URL. Any column in the published data can be referenced by enclosing the index number in curly braces “{}”. The taxon identifier in the core data file can also be referenced via the variable “[id].” The following examples illustrate these features:

1. The Integrated Taxonomic Information System (ITIS) uses Taxonomic Serial Numbers (TSN) to provide links to taxon pages on its web site.

http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=174375

²¹ <http://rs.gbif.org/extension/gbif/1.0/identifier.xml>

²² <http://www.iana.org/assignments/media-types/index.html>

²³ http://links.gbif.org/gbif_dwc-a_metafile_en_v1

If a core data file is published using the ITIS TSN system a link can be composed and tied to the “identifier” term in the core data standard using the following syntax.

a record id based link to the species page:

```
<field
default="http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_valu
e={id}" term="http://purl.org/dc/terms/identifier"/>
```

where the original numeric value is replaced by the variable “{id}”. This value would be derived from the core ID.

2. The 2010 Catalogue of Life Annual Checklist provides similar identifiers. It also supports name-based searches that can also be encoded as URLs. For example, <http://www.catalogueoflife.org/annual-checklist/2010/search/all/key/Struthio+camelus/match/1> embeds a scientific name “Struthio camelus” into a URL. Full scientific name combinations can be published in the core data file using the Darwin Core term “scientificName.” If we assume that this term represented the 12th column in our core data file we could use the syntax

a record id based link to the species page:

```
<field default="http://www.catalogueoflife.org/annual-
checklist/2010/search/all/key/{12}/match/1"
term="http://purl.org/dc/terms/identifier"/>
```

where {12} represents the 12th column value that will be substituted in the URL.

Alternative Identifiers extension

Use this extension to provide one or more links to source data records related to the core taxon. A source database, may, for example provide access to the source data records through a web page, a web service, and a resolvable identifier such as LSID, DOI or other means.

RowType: <http://rs.gbif.org/terms/1.0/Identifier>

Repository: <http://rs.gbif.org/extension/gbif/1.0/identifier.xml>

taxonID	<p>The identifier used in the core data file representing the taxon for which the current record refers. This identifier provides the link between the core data record and the extension record.</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/taxonID Data Type: string Required: true</p>
identifier	Other known identifier used for the same taxon. Can be a URL pointing to a webpage, an xml or rdf document, a DOI, UUID or any other identifier

	<p>Example: "urn:lsid:ipni.org:names:692570-1:1.4"</p> <p>Qualified Name: http://purl.org/dc/terms/identifier Data Type: string Required: true</p>
title	<p>An optional display label for the URL that the publisher may prefer be displayed with the identifier or link</p> <p>Example: "Danaus plexippus page", "COL Taxon LSID"</p> <p>Qualified Name: http://purl.org/dc/terms/title Data Type: string Required: false</p>
subject	<p>keywords qualifying the identifier</p> <p>Qualified Name: http://purl.org/dc/terms/subject Data Type: string Required: false</p>
format	<p>mime type of content returned by identifier in case the identifier is resolvable. Plain UUIDs for example do not have a dc:format return type, as they are not resolvable on their own. For a list of MIME types see the list maintained by IANA: http://www.iana.org/assignments/media-types/index.html, in particular the text http://www.iana.org/assignments/media-types/text/ and application http://www.iana.org/assignments/media-types/application/ types. Frequently used values are text/html, text/xml, application/rdf+xml, application/json</p> <p>Example: application/rdf+xml</p> <p>Qualified Name: http://purl.org/dc/terms/format Data Type: string Required: false</p>

Resource Relationship Extension

This extension is used to describe one or more relationships that the core taxon has with other taxa, either in the source list or included in the record. This extension could be used, for example, to provide a list of plant species (one record per species) pollinated by a bee species listed in the core species list.

RowType: <http://rs.tdwg.org/dwc/terms/ResourceRelationship>
 Repository: http://rs.gbif.org/extension/dwc/resource_relation.xml

taxonID	<p>The identifier used in the core data file representing the taxon for which the current record refers. This identifier provides the link between the core data record and the extension record.</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/taxonID Data Type: string Required: true</p>
resourceRelationshipID	<p>An identifier for an instance of relationship between one resource (the subject) and another (relatedResource, the object).</p> <p>Example: "231" "urn:lsid:gbif.org:usages:32567"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/resourceRelationshipID Data Type: string Required: false</p>
relatedResourceId	<p>An identifier for a related resource (the object, rather than the subject of the relationship).</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/relatedResourceId Data Type: string Required: false</p>
relationshipOfResource	<p>The relationship of the resource identified by relatedResourceId to the subject (optionally identified by the resourceId). Recommended best practice is to use a controlled vocabulary.</p> <p>Example: "duplicate of", "mother of", "endoparasite of", "host to", "sibling of", "valid synonym of", "located within"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/relationshipOfResource Data Type: string Required: false</p>
relationshipAccordingTo	<p>The source (person, organization, publication, reference) establishing the relationship between the two resources.</p> <p>Example: "Julie Woodruff"</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/relationshipAccordingTo Data Type: string Required: false</p>
relationshipEstablishedDate	<p>The date-time on which the relationship between the two resources was established. Recommended best practice is to use an encoding scheme, such as ISO 8601:2004(E).</p> <p>Example: 1963-03-08T14:07-0600</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/relationshipEstablishedDate Data Type: string Required: false</p>
relationshipRemarks	Comments or notes about the relationship between the two resources.

	<p>Example: “mother and offspring collected from the same nest”</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/relationshipRemarks</p> <p>Data Type: string</p> <p>Required: false</p>
scientificName	<p>A taxon name that is the object (relatedResource) of an association with the subject (either a taxon or a taxon occurrence)</p> <p>Example: “<i>Quercus agrifolia</i> var. <i>oxyadenia</i> (Torr.)”</p> <p>Qualified Name: http://rs.tdwg.org/dwc/terms/scientificName</p> <p>Data Type: string</p> <p>Required: false</p>

Summary: Abbreviated listing of the GNA Profile

An abbreviated listing of the core terms and extensions with their terms is listed below.

Taxon

taxonID | acceptedNameUsageID | acceptedNameUsage | parentNameUsageID | parentNameUsage | originalNameUsageID | originalNameUsage | nameAccordingToID | nameAccordingTo | namePublishedInID | namePublishedIn | taxonConceptID | scientificNameID | scientificName | scientificNameAuthorship | higherClassification | kingdom | phylum | class | order | family | genus | subgenus | specificEpithet | infraspecificEpithet | taxonRank | verbatimTaxonRank | vernacularName | nomenclaturalCode | taxonomicStatus | nomenclaturalStatus | taxonRemarks | modified | language | rights | rightsHolder | accessRights | bibliographicCitation | informationWithheld | datasetID | datasetName | source

Extensions

Vernacular Names Extension

RowType: <http://rs.gbif.org/terms/1.0/VernacularName>
 Repository: <http://rs.gbif.org/extension/gbif/1.0/vernacular.xml>

vernacularName | source | language | temporal | locationID | locality | countryCode | sex | lifeStage | isPlural | isPreferredName | organismPart | taxonRemarks

References Extension

RowType: <http://rs.gbif.org/terms/1.0/Reference>
 Repository: <http://rs.gbif.org/extension/gbif/1.0/references.xml>

identifier | bibliographicCitation | title | creator | date | source | description | subject | language | rights | taxonRemarks | type

Species Distribution Extension

RowType: <http://rs.gbif.org/terms/1.0/Distribution>
 Repository: <http://rs.gbif.org/extension/gbif/1.0/distribution.xml>

locationID | locality | countryCode | lifeStage | occurrenceStatus | threatStatus | establishmentMeans | appendixCITES | eventDate | startDayOfYear | endDayOfYear | source | occurrenceRemarks

Species Description Extension

RowType: <http://rs.gbif.org/terms/1.0/Description>
 Repository: <http://rs.gbif.org/extension/gbif/1.0/description.xml>

| description | type | source | language | creator | contributor | audience | license | rightsHolder

Alternative Identifiers

RowType: <http://rs.gbif.org/terms/1.0/Identifier>

Repository: <http://rs.gbif.org/extension/gbif/1.0/identifier.xml>

identifier | title | subject | format |

Types and Specimens Extension

RowType: <http://rs.gbif.org/terms/1.0/Specimen>

Repository: http://rs.gbif.org/sandbox/extension/gbif_taxon_type_specimens.xml

bibliographicCitation | occurrenceID | institutionCode | collectionCode | catalogNumber | typeStatus | scientificName | taxonRank | identificationRemarks | locality | sex | recordedBy | source | verbatimEventDate | verbatimLabel | verbatimLongitude | verbatimLatitude

Resource Relationship Extension

RowType: <http://rs.tdwg.org/dwc/terms/ResourceRelationship>

Repository: http://rs.gbif.org/extension/dwc/resource_relation.xml

resourceRelationshipID | relatedResourceID | relationshipOfResource | relationshipAccordingTo | relationshipEstablishedDate | relationshipRemarks | scientificName